A Corpus-Based Study of Nationality-Related Swearwords in British National Corpus

Ilyas Babar Awan[1,] Muhammad Hussain[2]

[1]Lecturer in English (PhD Scholar), Translation & Interpretation Department, National University of Modern Languages, Islamabad
[2]PhD English (Scholar), Faculty of Language & Literature, International Islamic University, Islamabad

Abstract

Swearwords are a part of everyday language use. A number of corpus-based studies such as gender, age, and social class have been conducted; however, nationality-related swearwords are not explored particularly with reference to British National Corpus (BNC). The aim of the present study is to investigate the use of selected swearwords that provide evidence for ethnic/national biased approach of one nationality for the other. The present study identifies and classifies selected swearwords by using BNC in order to find out occurrences of the target words, frequencies and collocations. It further traces out the differences between their use in spoken and written expressions. The results of the study show that British native speakers use swearword 'mick' against Irish nationals frequently which is 15.22 per million words and the frequency of 'Guido' which is used for Italian in BNC is 4.81 per million words. Furthermore, in BNC spoken expression the second highest number of swearword 'Paki' is against Pakistani nationals which is 1.06 per million words. Future researchers may further explore the BNC in different perspectives.

Keywords: corpus, nationality, swearwords, British National Corpus (BNC)

1. Introduction

The present paper explores the use of nationality-related racial slurs in terms of swearwords with reference to British National Corpus (BNC). The paper aims to investigate how various nationalities have been stereotyped by using frequent racial slurs against them in BNC. The study may be helpful to understand the western biased approach against nationalities across the world. The study may serve as pattern study to explore BNC in various perspectives that have received little attention of researchers. Besides, the article may also be a contribution in the domain of the empire writes back. Swearwords are the category in linguistic items which represent and channelize the anger, frustration, disgust, and hatred (Fägersten, 2012; Dewaele, 2015). By and large, it is a common characteristic of the world languages that they contain swearwords. Swearing words reflect the attitude of language users toward the opposite gender, class, relation, and nationality (Chen, 2004; Ljung, 2010). Some commonly used sex-oriented swearing words are not too old in the English language, we find their traces in medieval ages, and they are used in multiple connotations (McEnery, 2004; Montagu, 1967). Languages borrow lexical items from other languages as a natural process of cultural invasion; therefore, they also incorporate swearing words in them. With the changing socio-political scenarios, the use of nationality-based swearing words has been a contemporary phenomenon that needs to be further investigated in language. This study is a quantitative and systematic account of swearwords in the English language as found in BNC. British National Corpus (BNC) is a 100,000,000-word corpus that provides empirical evidence of the use of swearwords against particular nationalities as well as cultural and ethnical groups. The selected swearwords have been identified in multiple ways as collocation and concordance. By comparing certain swearwords, we have associated socio-

political and cultural perspectives with them. The aim of this study is to analyse the use of swearwords in written and spoken genre.

2. Literature Review

This part of the study is an extension of the previous research works in the area of swearwords in general and BNC in particular. This study is about the use of nationality-based swearwords and how they reflect the resistance and hatred against particular nationalities, as revealed in the British National Corpus. Besides, bulk of studies (see Popuşoi, Havârneanu, & Havârneanu, 2018; Zenner & Ruette, 2018; Harismayanti, Budasi, Lin & Ramendra, 2017; Arsana, Putra, & Ramendra, 2017; Khoirunnisa & Wijayanto, 2017; Stenström, 2017; Rullyanti, 2017; Marom, 2014; Kwon & Cho, 2016; Güvendir, 2015; Kurniawan, 2015; Fägersten, 2014; Christie, 2013; Praschinger, Pomikal, & Stieger, 2011; Gauthier, 2012; Fras, 2012; Kirk, 2013; Suyanto, 2010; Hjort, 2009; Dewaele, 2004, among others) have been conducted on swearwords in various perspectives in general; however, the BNC has not been explored in terms of swearwords so far.

2.1. What Constitutes a Swear Word?

Swearing words are also known as offensive or profane words. Language users' associate expletive attributes while decoding curse, hatred, anger, and bias (Goddard, 2015; Moore, 2102).

These swear words unfold the use of language in the following manners:

a) Through expletive attributes a speaker may indicate his irritation, anger, and occasionally approval.
b) Sometimes the use of foul words, such as *fucking* and *goddamn* are used to put emphasis on certain feeling or to showcase frustration.
c) Swearing words also reflect the change in age and youth's communication to deconstruct social interaction (Thorne, 2014).

2.2. Types of Swearing Word

Though the swearing words are considered bad language but no language of the world is swearing free (Darma, Wennyta, & Fitri, 2017). These words are categorized as:

a. Swear words related to religion (e.g. Jesus, heaven, hell and damn)
b. Swear words related to sex (e.g. f**k and c**t)
c. Swear words related to racism (e.g. nigger)
d. Swear words related to defecation (e.g. sh*t and pi**)
e. Swear words related to homophobia (e.g. queer)
f. Swear words related to nationality (e.g. paki, mick)

2.3. Functions and Motifs of Swearwords

Swearwords can be categorized in a number of ways and in accordance with different variables. One way of categorising swearwords is based on their motifs: religious, scatological, genital, intercourse, sodomy and mother (Ljung, 2006, p. 62). Another classification can be sex, religion, bodily functions, ethnic groups, food, dirt and death (Andersson & Trudgill, 1990, p. 55). Yet another order of grouping can be cursing, profanity, blasphemy, taboo, obscenity, vulgarity, slang, epithets, scatology, insults, and slurs (Jay, 1992, p. 1-8). A swearword, or phrase, can belong to more than one group at one and the same time like *fucking cunt*, which belongs to both the intercourse group and the genital group (Ljung, 2006, p. 63).

This study explores the use of swearwords in terms of nationality. It aims at finding out the frequencies of these words (Leech & Rayson, 2014) in overall corpus at first phase of the analysis which is followed by the second phase of analysis in which swearwords used in spoken corpus will be identified and their frequencies will also been mentioned. At the last phase of the analysis, the written texts will be analyzed in terms of swearwords used in order to express hatred for other nationalities.

3. Methodology

The aim of the present study is to explore the use of swearwords against nationalities with reference to BNC. For this purpose, various swearwords have been identified from various online sources (https://www.noswearing.com/dictionary). Having identified various swearwords, British National Corpus was used in order to find out frequent swearwords in terms of written and spoken perspectives. Overall swearwords were identified at the first phase of the analysis, followed by the identification of swearwords in the spoken and written corpus. Additionally, hits, total texts, words and frequencies were explored, identified, and classified in spoken and written data. The study employs quantitative and qualitative methods in order to investigate the selected swearwords by finding out the frequencies followed by interpretation of the data. Furthermore, corpus linguistics is itself termed as a fast-growing methodology in contemporary linguistics (Gries, 2009), thus frequencies and analysis have been carried out with the help of BNC results.

4. Analysis/Findings

The BNC data has been analyzed in this section. Our study has mainly focused on the words showing disrespect or hatred against a particular nationality. From BNC corpus, we have obtained the following results about the selected swearwords.

Table 1

Swearwords in Overall BNC Data

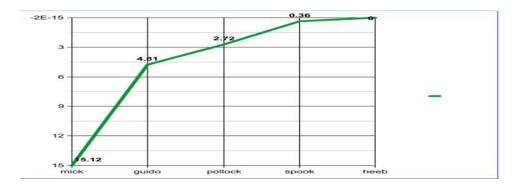| Swearwords	Nationality	Hits	Text/total Text	Words	Frequency
Mick	Irish	1496	300/40,48	98,313,429	15.22 instance per million word
Guido	Italian	473	39	Do	4.81
Pollock	Polish	267	95	Do	2.72
Spook	White	35	26	Do	0.36
Paki	Pakistani	28	18	Do	0.28
Spick	Mexican American	24	23	Do	0.24
Coon	African American	17	14	Do	0.17
Wop	Italian	14	10	Do	0.14
Wetback	Mexican	1	1	Do	0.01
Ruski	Russian	1	1	Do	0.01
Heeb	Jewish	0	0	0	0
Porch monkey	African American	0	0	0	0

Figure 1. Quantitative findings of selected swearwords form Overall BNC Data
https://corpus.byu.edu/bnc

Table 2

Swearwords in Spoken BNC Data

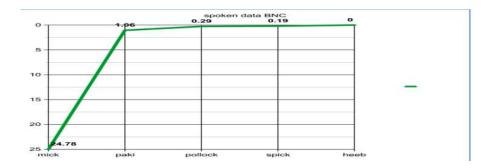| Swearwords	Hits	Words/total Words	Words	Frequency
Mick	258	63/908	10,409,858	24.78 instance per million word
Paki	11	5	Do	1.06
Pollock	3	1	Do	0.29
Spick	2	2	Do	0.19
Spook	2	2	Do	0.19
Guido	1	1	Do	0.1
Porch monkey	0	0	0	0
Wetback	0	0	0	0
Heeb	0	0	0	0
Ruski	0	0	0	0
Coon	0	0	0	0
Porch monkey	0	0	0	0

Figure 2. Quantitative findings of selected swearing words from spoken BNC data
https://corpus.byu.edu/bnc

Table 3

Swearwords in Written BNC Data

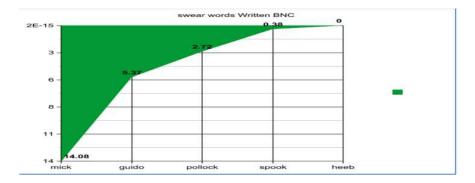| Swearwords	Hits	Text/total text	Words	Frequency
Mick	1238	237/3140	87,903,571	14.08 instance per million word
Guido	472	38	Do	5.37
Pollock	267	95	Do	2.72
Spook	33	24	Do	0.38
Spick	22	21	Do	0.25
Paki	17	13	Do	0.19
Coon	17	14	Do	0.19
Wop	10	9	Do	0.11
Wetback	1	1	Do	0.01
Ruski	1	1	Do	0.01
Heeb	0	0	0	0
Porch monkey	0	0	0	0

Figure 3. Quantitative findings of selected swearwords from written BNC data
https://corpus.byu.edu/bnc

5. Socio-Political Perspective Behind Swearwords

The purpose of collecting the data of nationality-based swearwords from BNC is to provide evidence against the British community that claims to be a multicultural community and pays respect to other cultural and ethnic groups. The result of Figure1 shows that British native speakers use most frequent swearword against Irish national (1496 hits from 300 texts). The frequency of the swearword *mick* (Irish) is 15.22 per million words. The second highest swearword used against Italian is *guido* (473 hits out of 39 texts). The frequency of *guido* in BNC is 4.81 per million words. The third highest swearword is used against Polish, which is *pollock* (267 hits out of 95 texts). The frequency of *pollock* in BNC is 2.72 per million words. Anglo-Irish relationships with respect to the swearword *mick* against Irish may be evaluated under the light of BNC. The Act of Union between these two cultural and national identities came into effect on January 1, 1801.

> That, from the first day of January one thousand eight hundred and one, all prohibitions and bounties on the exports of articles, the growth, produce, or manufacture of either country, to the other, shall cease…. That all articles, the growth, produce, or manufacture of either country… shall from thenceforth be imported into each country from the other, free from duty…
> (Act of Union, 1st August 1800; available at www.statutelaw.gov.uk)

Among ethnic identities living in Ireland, 90% are Irish people whereas in England the ratio of White British is 81.9%. About 700,000 Irish born people live in England. England's political relationships have never been healthy with Ireland. During the twentieth century in Northern Ireland, ethnonational conflict ignited ethnic fabric of both the nations. This conflict was primarily the political one, but later on it turned to be sectarian. Ireland's 84% population is Roman Catholic whereas England's 71.8% population is Protestant. Keeping in view the ethnic fragmentation of Anglo-Irish society, swearwords both in written and spoken expressions, show the disgrace, disrespect, and hatred against the opponent society, culture, and nationality. This is what BNC data show that the most swearwords have been used against Irish nationals. This is the highest ratio against Irish in BNC both in written and spoken language, respectively.

On the contrary, we do not find swearword in BNC for African Americans (Porch Monkey) and Jews (Heeb). One of the impressions for the absence of these swearwords may be on the credibility and representativeness of BNC. The political and social status of these two cultural and national identities has manifold cultural imprints on British society.

In BNC spoken corpus, the 2nd highest number of swearwords is against Pakistani nationals after Irish. BNC is a representative data from 1980 to 1993. Second highest number of swearwords against Pakistan reflects the socio-political relationship between both the nations. During the Zia-ul-Haq regime, many a sociopolitical changes took place in Pakistani society. On May 26, 1980 for the first time in Pakistan Federal Sharia Courts were established. In 1981, Interest-free banking was introduced, and on March 11, 1983 a nuclear test was carried out by Pakistan without announcing. Being a former colony of Great Britain such changes in indigenous policies and political narratives might not be expected from British, so they started to reflect their hatred particularly against Pakistani diaspora in spoken expression. This is evident form the BNC data that swearwords accommodate national interests of a nation.

6. Conclusion

After presenting the findings in detail, this section of the study offers a brief conclusion drawn from the study. Language as a tool defines the cultural and sociopolitical narratives in written (literary, journalitic, academic, etc.) and spoken text. The use of specific swearwords may provide evidence to reach the psychological and cultural changes that take place in a society. This study has provided us database evidence to deconstruct language and

society at the same time. The research questions posed in the introductory section have been addressed on the basis of findings from the BNC. To sum up, BNC has provided evidence about the selected swearwords that have been used by native speakers to channelize their invasive variants. Speakers, particularly in spoken variety of language, have been seen using highly aggressive tokens against some nationalities. This study also reveals that a number of nationality-based swearwords by native speakers reflect a psychological superiority complex and hatred against opposing nationalities. This has also been noticed that the conflicts of England with its neighboring nations have been reflected in written and spoken expression by the English. In addition, the paradigm shift in political and cultural narratives has also been reflected in written and spoken expressions in BNC. The use of swearwords by a particular society needs to be psychoanalysed in later studies.

References

Anderson, L.G., & Trudgill, P. (1990). *Bad language.* England: Penguin Books.

Arsana, M., Putra, I. N. A. J., & Ramendra, D. P. (2017). An analysis of balinese swear words among the northern balinese teenagers in seririt district in their daily conversation. *Jurnal Pendidikan Bahasa Inggris*, *3*(3).

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language Structure and use.* Cambridge, UK: Cambridge University Press.

British National Corpus (https://corpus.byu.edu/bnc).

Bruin, M. D. (2018). *Translating swear words: a fucking difficult task. An analysis of Sam Garrett's American English translation of Turks Fruit* (Master's thesis).

Chen, C. (2004). On the Hong Kong Chinese Subtitling of English Swearwords. *Meta: Journal des traducteurs/Meta: Translators' Journal, 49*(1), 135-147.

Christie, C. (2013). The relevance of taboo language: An analysis of the indexical values of swearwords. *Journal of Pragmatics*, *58*, 152-169.

Darma, R. S., Wennyta, W., & Fitri, N. (2017). Swearing Words in "Celebrity Read Mean Tweets" in Jimmy Kimmel Late Night Show: A Sociolinguistic Study. *JELT: Journal of English Language Teaching, 1*(1).

Dewaele, J. M. (2015). British 'Bollocks' versus American 'Jerk': Do native British English speakers swear more–or differently–compared to American English speakers? *Applied Linguistics Review, 6*(3), 309-339.

Dewaele, J. M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development*, *25*(2-3), 204-222.

Fägersten, K. B. (2014). he use of English swear words in Swedish media.

Fägersten, K. B. (2012). Who's swearing now? The social aspects of conversational swearing. Cambridge Scholars Publishing.

Fras, J. (2012). The ideology of swearwords in Slovenia. *Language & Communication, 32*(3), 229-239.

Gauthier, M. (2012). Profanity and gender: A diachronic analysis of men's and women's use and perception of swear words. *Profanity and Gender: A Diachronic Analysis of Men's and Women's Use and Perception of Swear Words.*

Goddard, C. (2015). "Swear words" and "curse words" in Australian (and American) English. At the crossroads of pragmatics, semantics and sociolinguistics. *Intercultural Pragmatics, 12*(2), 189-218.

Gries, S. T. (2009). What is corpus linguistics? *Language and linguistics compass, 3*(5), 1225-1241.

Gries, S. T. (2009). *Quantitative corpus linguistics with R: A practical introduction.* New York, NY: Routledge.

Güvendir, E. (2015). Why are males inclined to use strong swear words more than females? An evolutionary explanation based on male intergroup aggressiveness. *Language Sciences,* 50, 133-139.

Hjort, M. (2009). Swearwords in subtitles: A balancing act. *TRAlinea, Special Issue: The Translation of Dialects in Multimedia (2009).*

Harismayanti, I., Budasi, I. G., Lin, D. A., & Ramendra, D. P. (2017). An Analysis of Swear Words Used by the Teenagers in Singaraja. *Jurnal Pendidikan Bahasa Inggris undiksha, 4*(2).

Jay, T. (1992). Cursing in America: A Psycholinguistic Study of Dirty Language in the Courts, in the Movies, in the Schoolyards, and on the Streets. Philadelphia: John Benjamins.

Khoirunnisa, F. R., & Wijayanto, A. (2017). *The Use Of Swear Words In Deadpool Movie: A Sociopragmatic Study* (Doctoral dissertation, Universitas Muhammadiyah Surakarta).

Kirk, C. (2013). The most popular swear words on Facebook. *Slate.com.*

Kurniawan, K. E. (2015). *The analysis of taboo words and swear words in george carlin's monologue seven words you can't say on television* (Doctoral dissertation, SANATA DHARMA UNIVERSITY).

Kwon, K. H., & Cho, D. (2016). Reimagining swearing effects online: the use of swearwords in news audience commenting. *AoIR Selected Papers of Internet Research*, *5*.

Leech, G., & Rayson, P. (2014). *Word frequencies in written and spoken English: Based on the British National Corpus.* Routledge.

Ljung, M. (2010). *Swearing: A cross-cultural linguistic study.* Springer.

Ljung, M. (2006). *Svordomsboken.* Uddevalla: MediaPrint.

Marom, N. (2014). In Case of a Curse: *The Translation of Swearwords in Hebrew Subtitles of American Films and Television Series.* University of Haifa, Faculty of the Humanities, Department of English Language and Literature.

McEnery, T. (2004). *Swearing in English: Bad language, purity and power from 1586 to the present.* Routledge.

McEnery, A., & Xiao, Z. (2004). Swearing in modern British English: the case of fuck in the BNC. *Language and Literature*, *13*(3), 235-268.

Montagu, A. (1967). *The anatomy of swearing.* University of Pennsylvania press.

Moore, R. L. (2012). On swearwords and slang. *American speech, 87*(2), 170-189.

Praschinger, A., Pomikal, C., & Stieger, S. (2011). May I curse a referee? Swear words and consequences. *Journal of sports science & medicine, 10*(2), 341.

Popuşoi, S. A., Havârneanu, G. M., & Havârneanu, C. E. (2018). "Get the f#∗ k out of my way!" Exploring the cathartic effect of swearwords in coping with driving anger. *Transportation Research Part F: Traffic Psychology and Behaviour,* 56, 215-226.

Rullyanti, M. (2017). Swearwords used by Bengkulunese: A Case Study of Taboo Expressions. *Premise: Journal of English Education, 3*(1), 18-27.

Stenström, A. B. (2017). *Swearing in English and Spanish teenage talk. Advances in Swearing Research: New languages and new contexts,* 282, 157.

Suyanto, A. (2010). Javanese swear words in a boarding house: The case of five Soegijapranata Catholic university students. *Unpublished thesis. Soegijapranata Catholic University, Semarang.*

Thorne, T. (2014). *Dictionary of contemporary slang.* Bloomsbury Publishing.

Zenner, E., & Ruette, T. (2018). The borrowability of English swearwords in Dutch: a variationist approach.